Subdifferentials, Faces, and Dual Matrices

K. Ziętak

Institute of Computer Science
University of Wrocław
ul. Przesmyckiego 20
51-151 Wrocław, Poland

Submitted by George P. Barker

ABSTRACT

A characterization of the dual matrices for the unitarily invariant norms is given. Moreover, the connection between the dual matrices, the subdifferentials of matrix norms, and the faces of the unit ball of matrices is presented.

---

1. INTRODUCTION

Let a complex $m \times n$ matrix $A = [a_{ij}]$, $A \in \mathbb{C}^{m \times n}$, have the following singular value composition (SVD):

$$A = U \Sigma V^H, \tag{1.1}$$

where U and V are unitary matrices of order m and n, respectively, and Σ is a diagonal matrix in $\mathbb{R}^{m \times n}$ with singular values $\sigma_j = \sigma_j(A)$,

$$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_t \geqslant 0, \qquad t = \min\{m, n\},$$

on the diagonal. A norm $\|\cdot\|$ is unitarily invariant if $\|A\| = \|PAQ\|$ for arbitrary unitary matrices P and Q. Therefore $\|A\| = \|\Sigma\|$.

655 Avenue of the Americas, New York, NY 10010 0024-3795/93/$6.00

It has been shown by von Neumann [6] that any unitarily invariant norm $\|\cdot\|$ of a matrix is defined by the symmetric gauge function Φ (see, for example, [9, p. 64])

$$\|A\| = \Phi(\sigma(A)),$$

where

$$\sigma \equiv \sigma(A) = \left[\sigma_1(A), \dots, \sigma_t(A)\right]^T \in \mathbb{R}^t.$$

We use the following notation:

$$\operatorname{diag}(\sigma) \equiv \operatorname{diag}(\sigma(A)) \equiv \operatorname{diag}\big(\sigma_j(A)\big) = \Sigma.$$

The matrix norm, generated by the symmetric gauge function Φ, is denoted by $\|\cdot\|_\Phi$ where necessary. The symmetric gauge function Φ is a norm in $\mathbb{R}^t$, and its value does not depend on the order and signs of the components of a vector. The l_p-norm in $\mathbb{R}^t$, $1 \leqslant p \leqslant \infty$, is a symmetric gauge function. Unitarily invariant norms corresponding to l_p-norms are called Schatten (or c_p) norms, and we denote them by $\|\cdot\|_p$:

$$\|A\|_p = \|\sigma(A)\|_p.$$

For $p = \infty$ this is the spectral norm, and for $p = 1$ we obtain the nuclear (or trace) norm.

The polar Φ^* of the symmetric gauge function Φ, defined as

$$\Phi^*(x) = \max_{\Phi(y) \leqslant 1} y^T x,$$

is also a symmetric gauge function. Let

$$\langle A, X\rangle = \operatorname{trace}(AX^H) = \sum_{i,j} a_{ij}\bar{x}_{ij}.$$

The dual norm $\|\cdot\|^*$ to the norm $\|\cdot\|$,

$$\|A\|^* = \max_{\|X\| \leqslant 1} |\langle A, X\rangle|, \tag{1.2}$$

is equal to (see [9, p. 78])

$$\|A\|^* = \max_{\|X\| \leqslant 1} \operatorname{Re}\langle A, X \rangle. \tag{1.3}$$

If the norm $\|\cdot\|$ corresponds to the symmetric gauge function Φ then $\|A\|^* = \Phi^*(\sigma(A))$. Let

$$\frac{1}{p} + \frac{1}{q} = 1 \qquad (1 \leqslant p \leqslant \infty).$$

Then the c_q-norm is dual to the c_p-norm.

A matrix G for which the maximum (1.3) is reached is called a $\|\cdot\|$-dual matrix to A. We denote the set of $\|\cdot\|$-dual matrices G to A by $\mathbb{V}(A; \|\cdot\|)$. If $A \neq 0$ then $\|G\| = 1$. Hence we have for $A \neq 0$

$$\mathbb{V}(A: \|\cdot\|) = \{G : G \in \mathbb{C}^{m\times n}, \operatorname{Re}\langle A, G\rangle = \|A\|^*, \|G\| = 1\}. \tag{1.4}$$

In this paper we give a characterization of the set of $\|\cdot\|$-dual matrices for arbitrary unitarily invariant norms. This is a generalization of the results given for Schatten's norms and real matrices in [12] and is some improvement on Watson's characterization of the subdifferential of $\|A\|$ (see [10]). We also present some properties of the faces of the unit ball of matrices.

2. KNOWN RESULTS

The properties of $\|\cdot\|$-dual matrices were investigated in [12] for real rectangular matrices and unitarily invariant norms. In particular, it is known that for any arbitrary unitarily invariant norm $\|\cdot\|$, if G is a $\|\cdot\|$-dual matrix to a real rectangular matrix A, $A \neq 0$, then (see [12])

$$\langle \sigma(A), \sigma(G)\rangle = \Phi^*(\sigma(A)), \qquad \Phi(\sigma(G)) = 1. \tag{2.1}$$

The vector $\sigma(G)$ is the Φ-dual vector to $\sigma(A)$. It is easy to verify that the property (2.1) holds also for the complex matrices.

Now let A be a real matrix, $A \in \mathbb{R}^{m\times n}$. We recall the definition of the subdifferential (see [7]). We formulate this definition for a matrix norm. The subdifferential of $\|A\|$ is denoted by $\partial\|A\|$, and it is defined in the following way:

$$\partial\|A\| = \{G : G \in \mathbb{R}^{m\times n}, \|B\| \geqslant \|A\| + \langle B - A, G\rangle \text{ for all } B \in \mathbb{R}^{m\times n}\}.$$

Analogously we define $\partial\Phi(x)$ for $x \in \mathbb{R}^n$. It is well known that $G \in \partial\|A\|$ is equivalent to the statements (see for example [10])

$$\|A\| = \langle A, G\rangle, \qquad \|G\|^* \leqslant 1. \tag{2.2}$$

The roles of a norm and its dual can be interchanged in this definition.

V^0 det A be a real matrix. From (1.4) and (2.2) it follows that the subdifferential of $\|A\|^*$ is the set of $\|\cdot\|$-dual matrices to A:

$$\partial\|A\|^* = \mathbb{V}(A; \|\cdot\|).$$

Watson gave the following characterization of $\partial\|A\|_\Phi$ for any arbitrary unitarily invariant norm $\|\cdot\|_\Phi$. Let $\mathscr{S}(A)$ denote the following set of matrices:

$$\mathscr{S}(A) = \left\{G : G = U\operatorname{diag}(\tau_v)V^T \text{ is any SVD}, \; A = U\Sigma V^T, \; \tau \in \partial\Phi(\sigma(A))\right\}$$

In the definition of $\mathscr{S}(A)$ we take all SVD of A. Then (see [10])

$$\partial\|A\|_\Phi = \operatorname{conv}(\mathscr{S}(A)),$$

where conv denotes the convex hull of a set. However, for c_p-norms ($1 \leqslant p \leqslant \infty$), the formulae given in [12] imply that

$$\partial\|A\| = \mathscr{S}(A) \tag{2.3}$$

for the c_p-norms ($1 \leqslant p \leqslant \infty$).

In this paper we prove that (2.3) holds for any arbitrary unitarily invariant norm and arbitrary real matrix A.

In what follows we use the notation connected with the dual matrices instead of the subdifferentials. Moreover, without loss of generality we assume

$$m \geqslant n,$$

which simplifies notation and presentation.

3. CHARACTERIZATION OF DUAL MATRICES

Let a complex matrix A, $A \in \mathbb{C}^{m\times n}$, have the SVD (1.1), let $\|\cdot\|$ be any arbitrary unitarily invariant norm, and let $s = \operatorname{rank} A$. We now show that G is a $\|\cdot\|$-dual matrix to A if and only if G has the form

$$G = UDV^H, \qquad D \in \mathbb{V}(\Sigma; \|\cdot\|), \tag{3.1}$$

where Σ is determined as in (1.1). This is connected with the properties of the scalar product $\langle \cdot, \cdot \rangle$.

THEOREM 3.1. *Let $A \in \mathbb{C}^{m\times n}$ have the SVD (1.1). Then*

$$\mathbb{V}(A; \|\cdot\|) = \{G : G = UDV^H,\ D \in \mathbb{V}(\Sigma; \|\cdot\|)\}.$$

Proof. If $A = 0$, then the theorem is obvious. We now assume $A \neq 0$. Let G have the form (3.1). Then $\|D\| = 1$, so $\|G\| = 1$ and

$$\operatorname{Re}\langle A, G\rangle = \operatorname{Re}\langle \Sigma, D\rangle = \|\Sigma\|^*.$$

Therefore $G \in \mathbb{V}(A; \|\cdot\|)$. Now, let $G \in \mathbb{V}(A; \|\cdot\|)$. Then we have [see (2.1)]

$$\operatorname{diag}(\sigma_j(G)) \in \mathbb{V}(\Sigma; \|\cdot\|).$$

Thus

$$\operatorname{Re}\langle A, G\rangle = \operatorname{Re}\langle \Sigma, U^H GV\rangle = \|\Sigma\|^*$$

and $\|U^H GV\| = 1$. Hence

$$U^H GV \in \mathbb{V}(\Sigma; \|\cdot\|),$$

which completes the proof. ∎

Because of Theorem 3.1 it suffices to characterize the set of $\|\cdot\|$-dual matrices corresponding to a diagonal matrix $\Sigma = \operatorname{diag}(\sigma_j(A))$. We now prove some properties of matrices dual to Σ. First we recall the result of von Neumann (see, for example, [9, p. 76]):

$$\max_{X, Y \text{ unitary}} \operatorname{Re} \operatorname{trace}(AXBY^H) = \sum_j \sigma_j(A)\sigma_j(B). \tag{3.2}$$

Stewart and Ji-guang Sun gave a new proof of (3.2). From their proof we obtain the following corollary for diagonal matrices.

COROLLARY 3.1. *Let* $A = \operatorname{diag}(\sigma_j)$ *and* $B = \operatorname{diag}(\tau_j)$, $A, B \in \mathbb{R}^{n\times n}$,

$$\sigma_1 > \sigma_2 > \cdots > \sigma_n > 0, \tag{3.3}$$

$$\tau_1 \geqslant \tau_2 \geqslant \cdots \geqslant \tau_n \geqslant 0, \tag{3.4}$$

and let the maximum (3.2) *be attained for unitary matrices* X_0 *and* Y_0. *Then the matrix*

$$C = X_0 B Y_0^H \tag{3.5}$$

is diagonal with nonnegative diagonal elements.

We now extend Corollary 3.1 to rectangular diagonal matrices A and B, replacing the assumption (3.3) by

$$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \sigma_n \geqslant 0. \tag{3.6}$$

We introduce some auxiliary notation:

$$X_0 = [x_{ij}] = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}, \qquad Y_0 = [y_{ij}] = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix},$$

$$C = [c_{ij}] = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},$$

where $X_{11}, Y_{11} \in \mathbb{C}^{s\times r}$ and $C_{11} \in \mathbb{C}^{s\times s}$,

$$s = \operatorname{rank} A, \qquad r = \operatorname{rank} B. \tag{3.7}$$

LEMMA 3.1. *Let* $A = \operatorname{diag}(\sigma_j) \neq 0$ *and* $B = \operatorname{diag}(\tau_j) \neq 0$, $A, B \in \mathbb{R}^{m \times n}$, *where* σ_j *satisfy* (3.6) *and* τ_j *satisfy* (3.4). *If the assumptions* (3.7) *are satisfied, then* $X_{11} = Y_{11}$ *and*

$$c_{jj} \geqslant 0 \qquad (j = 1, \dots, s),$$

where X_0, Y_0, *and* C *are determined as in* (3.5). *In other words,* C_{11} *is a nonnegative definite matrix.*

Proof. Let the assumptions (3.4)–(3.7) be satisfied, and let the maximum (3.2) be attained for unitary matrices X_0 and Y_0. Then

$$\sum_j \sigma_j \tau_j = \operatorname{Re} \operatorname{trace}(\Sigma C) = \operatorname{Re} \sum_j \sigma_j c_{jj} = \operatorname{Re} \sum_j \sigma_j \sum_k x_{jk} \tau_k \bar{y}_{jk}.$$

It is easy to verify that

$$\operatorname{Re}\left(x_{jk}\bar{y}_{jk}\right) \leqslant \tfrac{1}{2}\left(|x_{jk}|^2 + |y_{jk}|^2\right). \tag{3.8}$$

Moreover, it is known that (see [3, 11]) if a matrix $H = [h_{ij}]$ satisfies

$$\begin{aligned} \sum_i h_{ij}^2 &\leqslant 1 \qquad \text{for every } j, \\ \sum_j h_{ij}^2 &\leqslant 1 \qquad \text{for every } i, \end{aligned} \tag{3.9}$$

then

$$\sum_{i,j} \sigma_i \tau_j h_{ij}^2 \leqslant \sum_j \sigma_j \tau_j. \tag{3.10}$$

The matrices X_0 and Y_0 are unitary. Therefore the matrix H with elements h_{ij} equal to $|x_{ij}|$ (or $|y_{ij}|$) satisfies the assumptions (3.9). Consequently, from (3.10) we have

$$\sum_{j,k} \sigma_j \tau_k |x_{jk}|^2 \leqslant \sum_j \sigma_j \tau_j,$$

$$\sum_{j,k} \sigma_j \tau_k |y_{jk}|^2 \leqslant \sum_j \sigma_j \tau_j.$$

Combining this with (3.8), we obtain

$$\sum_{j,k} \sigma_j \tau_k \operatorname{Re}\left(x_{jk} \bar{y}_{jk}\right) \leqslant \frac{1}{2} \sum_{j,k} \sigma_j \tau_k \left(|x_{jk}|^2 + |y_{jk}|^2\right) \leqslant \sum_j \sigma_j \tau_j.$$

This implies

$$\sum_{j=1}^{s} \sigma_j \sum_{k=1}^{r} \tau_k \left[|x_{jk}|^2 + |y_{jk}|^2 - 2 \operatorname{Re}\left(x_{jk} \bar{y}_{jk}\right)\right] = 0.$$

From this we obtain [see (3.8)]

$$2 \operatorname{Re}\left(x_{jk} \bar{y}_{jk}\right) = |x_{jk}|^2 + |y_{jk}|^2 \qquad (j = 1, \dots, s; \quad k = 1, \dots, r).$$

By the properties of complex numbers we have

$$x_{jk} = y_{jk} \qquad (j = 1, \dots, s; \quad k = 1, \dots, r),$$

so $X_{11} = Y_{11}$.

From the definition of C we obtain

$$c_{jj} = \sum_{k=1}^{r} x_{jk} \tau_k \bar{y}_{jk}.$$

Thus we have

$$c_{jj} = \sum_{k=1}^{r} \tau_k |x_{jk}|^2 \geqslant 0 \qquad (j = 1, \dots, s).$$

Moreover, C_{11} is nonnegative definite, because

$$C_{11} = X_{11} \operatorname{diag}(\tau_1, \dots, \tau_r)\, X_{11}^H,$$

which completes the proof. ∎

We now prove that if the assumptions of Lemma 3.1 are satisfied then $C_{21} = 0$ and $C_{12} = 0$.

LEMMA 3.2. *Let the assumptions of Lemma* 3.1 *be satisfied. Then* $C_{21} = 0$ *and* $C_{12} = 0$. *Moreover, if* $\sigma_j \neq \sigma_k$ *then* $c_{jk} = c_{kj} = 0$.

Proof. This proof is based on some ideas from the proof of Lemma 3.4 in [9]. However, we include it because the assumption (3.3) is replaced by (3.6) and we consider rectangular matrices.

Let us suppose *a contrario* that $C_{21} \neq 0$. Then there exist indices j and k, $s+1 \leqslant j \leqslant m$ and $1 \leqslant k \leqslant s$, such that $c_{jk} \neq 0$. We recall that $s = \operatorname{rank} A$. We multiply the jth row of C by $-\bar{c}_{jk}/|c_{jk}|$. If $j \leqslant n$, then we additionally divide the jth column by the same number. We denote the matrix obtained after these unitary transformations by $\tilde{C} = [\tilde{c}_{ij}]$. Of course, $\|\tilde{C}\| = \|C\|$ and the diagonal elements of $\tilde{C}$ are the same as of C, so

$$\operatorname{Re}\operatorname{trace}(\Sigma C) = \operatorname{Re}\operatorname{trace}(\Sigma \tilde{C}).$$

We have $\tilde{c}_{jk} < 0$.

Let $R_\theta = [r_{ij}] \in \mathbb{R}^{m\times m}$ denote a rotation in the (k, j) plane. The matrix R_θ differs from the identity matrix by four elements only:

$$r_{kk} = r_{jj} = \cos\theta, \qquad r_{jk} = -r_{kj} = \sin\theta. \tag{3.11}$$

We denote

$$\tilde{C}_\theta = R_\theta \tilde{C}. \tag{3.12}$$

If $j \leqslant n$ then

$$\operatorname{Re}\operatorname{trace}\left(\Sigma\tilde{C}_\theta\right) = \sigma_k\left(\cos\theta\, c_{kk} - \sin\theta\, \tilde{c}_{jk}\right) + \sigma_j \operatorname{Re}\left(\sin\theta\, \tilde{c}_{kj} + \cos\theta\, c_{jj}\right) + \sum_{i\neq j,k} \sigma_i c_{ii}. \tag{3.13}$$

If $j > n$ then

$$\operatorname{Re}\operatorname{trace}\left(\Sigma\tilde{C}_\theta\right) = \sigma_k\left(\cos\theta\, c_{kk} - \sin\theta\, \tilde{c}_{jk}\right) + \sum_{i\neq k} \sigma_i c_{ii}. \tag{3.14}$$

Because $s+1 \leqslant j$, we have $\sigma_j = 0$ in (3.13). From (3.13) and (3.14) we obtain

$$\left.\frac{d\, \mathrm{Re}\, \mathrm{trace}\left(\Sigma \tilde{C}_\theta\right)}{d\theta}\right|_{\theta=0} = -\sigma_k \tilde{c}_{jk} > 0.$$

Thus a small change in θ increases Re trace$(\Sigma \tilde{C}_\theta)$, which is a contradiction of the optimality of C. Hence $C_{21} = 0$.

Now, let us suppose *a contrario* that $C_{12} \neq 0$. Then there exist j and k, $1 \leqslant k \leqslant s$ and $s+1 \leqslant j \leqslant n$, such that $c_{kj} \neq 0$. By multiplying the kth row of C by $\bar{c}_{kj}/|c_{kj}|$ and dividing the kth column by the same number, we obtain $\tilde{C}$ such that $\tilde{c}_{kj} > 0$. Let $R_\theta \in \mathbb{R}^{n \times n}$ be determined as in (3.11). Let us consider the matrix

$$\hat{C}_\theta = C\tilde{R}_\theta.$$

Then

$$\left.\frac{d\, \mathrm{Re}\, \mathrm{trace}\left(\Sigma \hat{C}_\theta\right)}{d\theta}\right|_{\theta=0} = \sigma_k \tilde{c}_{kj} > 0.$$

By the same reasoning as before, this leads to a contradiction. Therefore $C_{12} = 0$. Now, let $1 \leqslant k < j \leqslant s$ and $\sigma_j \neq \sigma_k$. Then it is possible to show that $c_{kj} = c_{jk} = 0$. We omit the proof, because it is exactly the same as part of the proof of Lemma 3.4 in [9]. This completes the proof. ■

Let Σ be determined as in (1.1). We have not assumed in (1.1) that the singular values $\sigma_j = \sigma_j(A)$ are distinct. Therefore the vector σ can have the following form:

$$[\sigma_1, \ldots, \sigma_n]^T = [\mu_1, \ldots, \mu_1, \mu_2, \ldots, \mu_2, \ldots, \mu_q, \ldots, \mu_q, 0, \ldots, 0]^T, \tag{3.15}$$

where

$$\mu_1 > \mu_2 > \cdots > \mu_q > \mu_{q+1} = 0. \tag{3.16}$$

Let t_j be the multiplicity of μ_j, so for example

$$\sigma_1 = \sigma_2 = \cdots = \sigma_{t_1} = \mu_1 > \sigma_{t_1+1}.$$

It is easy to verify that every SVD of Σ has the form (see, for example, [2])

$$\Sigma = W\Sigma T^H, \tag{3.17}$$

where W and T have the following diagonal block form:

$$W = \operatorname{diag}(W_1, \dots, W_q, W_0), \qquad T = \operatorname{diag}(W_1, \dots, W_q, T_0) \tag{3.18}$$

with $W_0, T_0, W_1, \dots, W_q$ being arbitrary unitary matrices

$$W_j \in \mathbb{C}^{t_j \times t_j} \qquad (j = 1, \dots, q),$$

$$W_0 \in \mathbb{C}^{(m-m+t_{1+1})\times(m-n+t_{g+1})}, T_0 \in \mathbb{C}^{t_{q+1}\times t_{q+1}}.$$

Let $\|\cdot\|$ be any arbitrary unitary invariant norm, and let D be a $\|\cdot\|$-dual matrix to Σ. Then [see (2.1) and (3.1)]

$$\begin{aligned}\operatorname{Re}\langle\Sigma, D\rangle = \sum_j \sigma_j\sigma_j(D) &= \max_{X, Y \text{ unitary}} \operatorname{Re}\operatorname{trace}\left[\Sigma X \operatorname{diag}(\sigma(D))\, Y^H\right] \\ &= \max_{X, Y \text{ unitary}} \operatorname{Re}\left\langle \Sigma, Y \operatorname{diag}(\sigma(D))\, X^H\right\rangle.\end{aligned}$$

Therefore $D = [d_{ij}]$ has to have the form [see (3.2), (3.5), and Lemma 3.1]

$$D = \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix},$$

where D_{11} is nonnegative definite, $D_{11} \in \mathbb{R}^{s\times s}$, $s = \operatorname{rank}\Sigma$, and $d_{ij} = 0$ if $\sigma_i \neq \sigma_j$. Let (3.16) hold. Then D_{11} has to have the following diagonal block form:

$$D_{11} = \operatorname{diag}(G_1, G_2, \dots, G_q),$$

where

$$G_j \in \mathbb{C}^{t_j \times t_j}, \qquad G_j \text{ nonnegative definite}.$$

Therefore G_j has the following SVD:

$$G_j = U_j \operatorname{diag}(\sigma(G_j))\, U_j^H,$$

where U_j is unitary. The block D_{22} of D corresponds to $\mu_{q+1} = 0$ [see (3.16)]. Let D_{22} have the SVD

$$D_{22} = U_0 \operatorname{diag}(\sigma(D_{22}))\, V_0^H.$$

We denote

$$\beta = \begin{bmatrix} \sigma(G_1) \\ \vdots \\ \sigma(G_q) \\ \sigma(D_{22}) \end{bmatrix}.$$

Then $\beta \in \mathbb{R}^n$, and we can write D in the form

$$D = \tilde{W} \operatorname{diag}(\beta)\, \tilde{T}^H, \tag{3.19}$$

where

$$\tilde{W} = \operatorname{diag}(U_1, \dots, U_q, U_0),$$

$$\tilde{T} = \operatorname{diag}(U_1, \dots, U_q, V_0).$$

The unitary matrices $\tilde{W}$ and $\tilde{T}$ have the form (3.18). Therefore, they can be chosen in the SVD of Σ [see (3.17)]. Now we can formulate the main result.

THEOREM 3.2. *Let $\|\cdot\|$ be any arbitrary unitarily invariant norm corresponding to the symmetric gauge function Φ, and let Σ, determined as in* (1.1), *satisfy* (3.15) *and* (3.16). *Then*

$$\mathbb{V}(\Sigma; \|\cdot\|) = \{D : D = W \operatorname{diag}(\tau)\, T^H, \textit{ where } \Sigma = W \Sigma T^H, \text{ is any SVD at } \Sigma$$
$$\langle \sigma, \tau \rangle = \Phi^*(\sigma), \Phi(\tau) = 1\}. \tag{3.20}$$

REMARK. *In* (3.20) *we take all SVD of* Σ.

Proof. Let D belong to the right side of (3.20). Then [see (3.17), (3.18)]

$$\operatorname{Re}\langle \Sigma, D\rangle = \langle \sigma, \tau \rangle = \Phi^*(\sigma) \quad \text{and} \quad \|D\| = 1.$$

Therefore D is a $\|\cdot\|$-dual matrix to Σ.

Now, let D be a $\|\cdot\|$-dual matrix to Σ. Then D has the form (3.19). The unitary matrices in (3.19) are the unitary matrices from the SVD of Σ [see (3.17)]. Therefore we have

$$\langle \Sigma, D\rangle = \Phi^*(\sigma) = \langle \sigma, \beta \rangle, \qquad \Phi(\beta) = \|D\| = 1.$$

Hence D belongs to the right side of (3.20), which completes the proof. ■

From Theorems 3.1 and 3.2 we obtain the following corollary:

COROLLARY 3.2. *Let $\|\cdot\|$ be any arbitrary unitarily invariant norm, and let $A \in \mathbb{R}^{m\times n}$. Then the relation* (2.3) *holds.*

4. FACES

A convex subset $\mathbb{F}$ of the unit ball

$$\mathbb{E}^{m\times n} = \{E : E \in \mathbb{C}^{m\times n}, \|E\| \leqslant 1\}$$

is called a face of $\mathbb{E}^{m\times n}$ if $B, C \in \mathbb{E}^{m\times n}$, and $\lambda B + (1 - \lambda)C \in \mathbb{F}$ for some $0 < \lambda < 1$ implies $B, C \in \mathbb{F}$. The properties of the faces of the unit ball endowed with the dual norm to the operator norm were investigated for example by Grząślewicz [4] (compare [2]). This case corresponds to the trace norm. We now consider any arbitrary unitarily invariant norm.

The extreme points of $\mathbb{E}^{m\times n}$ are actually faces consisting of only one point. In other words, $G \in \mathbb{E}^{m\times n}$ is an extreme point of $\mathbb{E}^{m\times n}$ if and only if the relation

$$G = \tfrac{1}{2}(G_1 + G_2), \qquad G_1, G_2 \in \mathbb{E}^{m\times n},$$

implies $G = G_1 = G_2$. A characterization of extreme points of the unit sphere of real $m \times n$ matrices endowed with any arbitrary unitarily invariant norm is given by Zię̨tak in [12]. Her result is extended by So [8] for the case of complex $n \times n$ matrices. For the spectral norm and real matrices this characterization was obtained by Lau and Riha [5]. All these characterizations of extreme points essentially follow from the paper [1] by Arazy. Namely, Arazy gave a characterization of the extreme points of the unit ball of unitary matrix spaces of compact operators. Ziętak obtained her result independently of Arazy.

It is easy to verify that $\mathbb{V}(A;\ \|\cdot\|)$ is a face of $\mathbb{E}^{m\times n}$. Namely, let $G_1, G_2 \in \mathbb{E}^{m\times n}$, $\lambda G_1 + (1-\lambda)G_2 \in \mathbb{V}(A;\ \|\cdot\|)$ for some λ, $0<\lambda<1$. Then we have

$$\|A\|^* = \operatorname{Re}\langle A, \lambda G_1 + (1-\lambda)G_2\rangle = \lambda \operatorname{Re}\langle A, G_1\rangle$$
$$+(1-\lambda)\operatorname{Re}\langle A, G_2\rangle \leqslant \|A\|^* \tag{4.1}$$

and

$$|\operatorname{Re}\langle A, G_i\rangle| \leqslant |\langle A, G_i\rangle| \leqslant \|A\|^*\|G_i\| = \|A\|^* \qquad (i=1,2).$$

Therefore

$$\operatorname{Re}\langle A, G_1\rangle = \operatorname{Re}\langle A, G_2\rangle = \|A\|^*,$$

and consequently $G_1, G_2 \in \mathbb{V}(A;\ \|\cdot\|)$.

We now show that if $\mathbb{F}$ is a face of $\mathbb{E}^{m\times n}$ then there exists a matrix A such that $\mathbb{F} \subseteq \mathbb{V}(A;\ \|\cdot\|)$. If $\mathbb{F} = \mathbb{E}^{m\times n}$ then $A=0$, so this case is trival. Therefore we formulate the theorem for proper faces of $\mathbb{E}^{m\times n}$.

THEOREM 4.1. *Let a convex subset $\mathbb{F}$ of $\mathbb{E}^{m\times n}$ be a proper face. Then there exists a matrix A, $A \neq 0$, such that*

$$\mathbb{F} \subseteq \mathbb{V}(A; \|\cdot\|). \tag{4.2}$$

Proof. Let $\mathbb{F}$ be a proper face of $\mathbb{E}^{m\times n}$, and let $H \in \mathbb{F}$. Then $\|H\| = 1$. We consider the set $\mathbb{V}(H;\ \|\cdot\|^*)$. Hence [see (1.4)]

$$\mathbb{V}(H; \|\cdot\|^*) = \{G : G \in \mathbb{C}^{m\times n}, \operatorname{Re}\langle H, G\rangle = 1, \|G\|^* = 1\}. \tag{4.3}$$

We now show that

$$\bigcap_{H\in\mathbb{F}} \mathbb{V}(H; \|\cdot\|^*) \neq \varnothing.$$

The set $\mathbb{V}(H; \|\cdot\|^*\|)$ is bounded and complete, so it is compact. Therefore it is sufficient to verify that

$$\{\mathbb{V}(H; \|\cdot\|) : H \in \mathbb{F}\}$$

has the finite intersection property. Let $H_1, \dots, H_k \in \mathbb{F}$. Then $\|H_j\| = 1$, since $\mathbb{F}$ is proper. We take

$$\tilde{H} = \frac{1}{k}(H_1 + \cdots + H_k).$$

By the convexity of $\mathbb{F}$, we have $\tilde{H} \in \mathbb{F}$. Let us consider the set of $\|\cdot\|^*$-dual matrices to $\tilde{H}$ [see (4.3)]. This set is not empty. Let

$$G' \in \mathbb{V}(\tilde{H}; \|\cdot\|^*). \tag{4.4}$$

Then we have $\|G'\|^* = 1$ and [see (1.2) and (1.3)]

$$\begin{aligned} |\langle H_j, G'\rangle| &\leqslant \|H_j\|\,\|G'\|^* = 1, \\ \operatorname{Re}\langle H_j, G'\rangle &\leqslant \|H_j\|\,\|G'\|^* = 1. \end{aligned} \tag{4.5}$$

Moreover, from (4.4) we obtain

$$1 = \operatorname{Re}\langle \tilde{H}, G'\rangle = \frac{1}{k}\sum_{j=1}^{k} \operatorname{Re}\langle H_j, G'\rangle \leqslant 1.$$

From this and (4.5) it follows that

$$\operatorname{Re}\langle H_j, G'\rangle = 1 \qquad (j = 1, \dots, k).$$

This means that

$$G' \in \bigcap_{j=1}^{k} \mathbb{V}(H_j; \|\cdot\|^*).$$

Therefore there exists an A such that $\|A\|^* = 1$ and

$$A \in \bigcap_{H \in \mathbb{F}} \mathbb{V}(H; \|\cdot\|^*),$$

which means that we have

$$\operatorname{Re}\langle H, A\rangle = 1 = \|A\|^* \qquad \text{for every} \quad H \in \mathbb{F}. \tag{4.6}$$

We recall that $\|H\| = 1$. Therefore from (4.6) we obtain (4.2), which completes the proof. ■

In the general case the matrix A for which (4.2) holds is not determined uniquely. So [8] gave the formulae characterizing the faces $\mathbb{F}$ of the unit ball of matrices endowed with the spectral and nuclear norms. These formulae are exactly the same as the formulae characterizing the dual matrices with respect to the spectral and nuclear norms. Therefore, for the spectral and nuclear norms there always exists a matrix A such that

$$\mathbb{F} = \mathbb{V}(A; \|\cdot\|). \tag{4.7}$$

It would be interesting to give a characterization of unitarily invariant norms for which the relation (4.7) holds. We are going to continue our investigations.

We wish to thank Professor G. A. Watson from the University of Dundee for calling our attention to the paper [8], and to Professor R. Grząślewicz from the Technical University of Wrocław for reference [1].

The work presented in this paper was begun while the author was visiting the Department of Computer Science of the University of Manchester. She expresses here her thanks for organizing the stay and for financial support.

REFERENCES

1 J. Arazy, On the geometry of the unit ball of unitary matrix spaces, *Integral Equations Operator Theory* 4(2):151–171 (1981).

2 H. Berens and M. Finzel, A continuous selection of the metric projection, in *Numerical Methods of Approximation Theory*, Vol. 8, Birkhäuser, Basel, 1987, pp. 21–29.

3 K. Fan, Maximum properties and inequalities for the eigenvalues of completely continuous operators, *Proc. Nat. Acad. Sci. U.S.A.* 37:760–766 (1951).

4 R. Grząślewicz, Faces in the unit ball of the dual of $\mathscr{L}(R^n)$, *Math. Ann.* 270:535–540 (1985).

5 K. K. Lau and W. O. Riha, Characterization of best approximations in normed linear spaces of matrices by elements of finite-dimensional subspaces, *Linear Algebra Appl.* 35:109–120 (1981).
6 J. von Neumann, Some matrix inequalities and metrization of metric spaces, *Tomsk Univ. Rev.* 1:286–300 (1937); in *Collected Works*, Pergamon, Oxford, 1962, pp. 205–219.
7 R. T. Rockafellar, *Convex Analysis*, Princeton U.P., Princeton, N.J., 1970.
8 W. So, Facial structures of Schatten p-norms, *Linear and Multilinear Algebra* 27:207–212 (1990).
9 G. W. Stewart and Ji-guang Sun, *Matrix Perturbation Theory*, Academic, San Diego, 1990.
10 G. A. Watson, Characterization of the subdifferential of some matrix norms, *Linear Algebra Appl.*, 170:33–45 (1992).
11 R. Wijsmann, Constructing all smallest simultaneous confidence sets in a given class, with applications to MANOVA, *Ann. Statist.* 7:1003–1018 (1979).
12 K. Ziętak, On the characterization of the extremal points of the unit sphere of matrices, *Linear Algebra Appl.* 106:57–75 (1988).

Received 12 March 1992; final manuscript accepted 7 May 1992